(19)日本国特許庁(JP)

(12) 公開特許公報(A)

(11)特許出願公開番号

特開平7-298879

(43)公開日 平成7年(1995)11月14日

(51) Int.Cl.⁶

識別記号

庁内整理番号

FΙ

技術表示箇所

C 1 2 N 9/12

1/21

8828-4B

15/09

学会大会講演要旨集」に発表

ZNA

(C 1 2 N 9/12

9281 - 4B

C 1 2 N 15/00

ZNA A

審査請求 未請求 請求項の数16 OL (全 20 頁) 最終頁に続く

(21)出願番号

(22)出願日

特蘭平6-95109

特許法第30条第1項適用申請有り 平成5年11月10日、

社団法人日本生物工学会発行の「平成5年度日本生物工

平成6年(1994)5月9日

(71)出願人 000003160

東洋紡績株式会社

大阪府大阪市北区堂島浜2丁目2番8号

(72)発明者 今中 忠行

大阪府吹田市藤白台2-28-11

(72)発明者 髙木 昌宏

大阪府吹田市青山台1-3 C-58-207

(72)発明者 森川 正章

大阪府箕面市小野原東5丁目4-12-406

(72) 発明者 柿原 博文

滋賀県草津市東矢倉2-19-16

(54) 【発明の名称】 超好熱始原菌由来のDNAポリメラーゼ遺伝子およびその用途

(57)【要約】

【目的】 新規な耐熱性DNAポリメラーゼを提供す る。

【構成】 超好熱始原菌であるKOD1から耐熱性DN Aポリメラーゼをコードする遺伝子をクローニングし、 さらに大腸菌にて発現可能な遺伝子を得て、T7プロモ ーターで誘導可能なプラスミドベクターに挿入し、該プ ラスミドベクターで大腸菌を形質転換する耐熱性DNA ポリメラーゼの製造法および精製法。

【特許請求の範囲】

【請求項1】 超好熱始原菌KOD1由来のDNAポリメラーゼ。

【請求項2】 分子量が約86~92Kdaであることを特徴とする請求項1記載のDNAポリメラーゼ。

【請求項3】 組換え宿主細胞を用いて生産されたことを特徴とする請求項1記載のDNAポリメラーゼ。

【請求項4】 配列番号2に記載されるアミノ酸配列を 含有することを特徴とする請求項1記載のDNAポリメ ラーゼ。

【請求項5】 超好熱始原菌KOD1由来のDNAポリメラーゼをコードする単離されたDNA。

【請求項6】 配列番号2に記載されるアミノ酸配列をコードする塩基配列を含有することを特徴とする請求項5に記載される単離されたDNA。

【請求項7】 配列番号3に記載される塩基配列または その一部分を含有することを特徴とする請求項5に記載 される単離されたDNA。

【請求項8】 請求項5に記載されたDNAをベクター に挿入したDNA組換え発現ベクター。

【請求項9】 ベクターがpET-8c由来のベクターであることを特徴とする請求項8記載のDNA組換え発現ベクター(pET-pol)。

【請求項10】 請求項8に記載されるDNA組換え発現DNAベクターを用いて形質転換された組換え宿主細胞。

【請求項11】 宿主細胞が大腸菌であることを特徴とする請求項8記載の組換え宿主細胞。

【請求項12】 請求項10に記載される組換え宿主細胞を培養し、培養物からDNAポリメラーゼを採取することを特徴とする超好熱始原菌KOD1由来のDNAポリメラーゼの製造法。

【請求項13】 請求項10に記載される組換え宿主細胞を培養し、(a) 該組換え宿主細胞を集めた後、破砕し、細胞抽出物を調製し、(b) 組換え宿主細胞由来の不純蛋白質を除去する工程を含むことを特徴とする超好熱始原菌KOD1由来DNAポリメラーゼを精製する方法。

【請求項14】 組換え宿主細胞を破砕する方法が、超音波処理であることを特徴とする請求項13記載の超好 40熱始原菌KOD1由来DNAポリメラーゼを精製する方法。

【請求項15】 組換え宿主細胞由来の不純蛋白質を除去する工程が高温熱処理であることを特徴とする請求項13記載の超好熱始原菌KOD1由来DNAポリメラーゼを精製する方法。

【請求項16】 高温熱処理条件が、70℃以上、好ましくは90℃以上であることを特徴とする請求項15記載の超好熱始原菌KOD1由来DNAポリメラーゼを精製する方法。

【発明の詳細な説明】

[0001]

【産業上の利用分野】本発明は新規な超好熱始原菌KOD1由来のDNAポリメラーゼおよび該ポリメラーゼをコードする遺伝子ならびに該遺伝子を使用するDNAポリメラーゼの製造法に関する。

[0002]

【従来の技術】従来から大腸菌のような中温性細菌由来のDNAポリメラーゼおよび中温性細菌に感染するファージ由来のDNAポリメラーゼに関しては、既に多くの研究がなされている。また最近、ポリメラーゼ連鎖反応(PCR)等の核酸増幅を用いる組換えDNA技術に有用な耐熱性DNAポリメラーゼに関する研究も多くなされている。PCR反応に用いられる耐熱性DNAポリメラーゼとしては、主としてサーマス・サーモフィラス(Thermus thermophilus)由来のDNAポリメラーゼ(Tthポリメラーゼ)や、サーマス・アクアチカス(Thermus aquaticus)由来のDNAポリメラーゼ(などが用いられてきた。

[0003]

【発明が解決しようとする課題】しかしながら、従来知 られている耐熱性DNAポリメラーゼには、耐熱性を有 するものの、その熱安定性や、有機溶媒に対する安定性 に若干、問題を残している。また、核酸の取り込みの際 の正確性にも欠ける点があり、DNA配列決定やポリメ ラーゼ連鎖反応にこれらの酵素を用いるに当たり、解決 すべき課題が残っている。そのため、これらの欠点を解 消する新規な耐熱性DNAポリメラーゼが待ち望まれて いた。またピロコッカス・フリオサス(Pyrococcus furi osus) 由来の耐熱性DNAポリメラーゼ(Pfuポリメラー ゼ、〒092/09689、特開平5-328969号公報)、サーモコッ カス・リトラリス(Thermococcus litoralis)由来の耐熱 性DNAポリメラーゼ(Tliポリメラーゼ、特別平6-7160 号公報)なども知られている。しかしながら、これらの 熱安定性DNAポリメラーゼは、核酸の取り込みの際の 正確性はTagDNAポリメラーゼやTheDNAポリ メラーゼに比べ優れているが、完全なものではなく新規 な耐熱性DNAポリメラーゼが望まれていた。

[0004]

7 【課題を解決するための手段】本発明者らは熱安定性D NAポリメラーゼを生産する新規な超好熱始原菌の1種 を得ることに成功し、さらにその遺伝子を解明して、本 発明に到達した。すなわち本発明は超好熱始原菌KOD 1由来のDNAポリメラーゼである。

【0005】また本発明は超好熱始原菌KOD1由来のDNAポリメラーゼをコードする単離されたDNAである。

【0006】さらに本発明は超好熱始原菌KOD1由来 のDNAポリメラーゼをコードする単離されたDNAを 50 ベクターに挿入したDNA網換え発現ベクターである。

【0007】また本発明は超好熱始原菌KOD1由来の DNAポリメラーゼをコードする単離されたDNAをベ クターに挿入したDNA組換え発現ベクターを用いて形 質転換された組換え宿主細胞である。

【0008】本発明は超好熱始原菌KOD1由来のDN Aポリメラーゼをコードする単離されたDNAをベクタ 一に挿入したDNA組換え発現ベクターを用いて形質転 換された組換え宿主細胞を培養し、培養物から培養物か らDNAポリメラーゼを採取することを特徴とする超好 熱始原菌KOD1由来のDNAポリメラーゼの製造法で 10

【0009】また本発明は超好熱始原菌KOD1由来の*

細胞形態

生育温度範囲

最適生育温度

生育 p H範囲

最適pH

最適塩濃度

栄養要求性

酸紫要求性

細胞膜胎質

DNAのGC含量

【0011】超好熱始原菌KOD1株は、直径約1μm の球菌であり、複数の極鞭毛を有していた。この菌株は 菌学的性質からPfuDNAポリメラーゼ生産菌(Pyroc occus furiosus) およびTli (Vent) DNAポリ メラーゼ生産菌(Thermococcus litoralis)との菌縁関係 が示唆された。

【0012】本発明の耐熱性DNAポリメラーゼ遺伝子 のクローニングは、以下の方法により行う。クローニン 30 グの方法は、PfuDNAポリメラーゼの保存領域アミ ノ酸配列(Nucleic Acids Research, 1993, vol.21, No. 2, 259-265) に基づき、プライマーを設計し、合成す

【0013】まず超好熱始原菌KOD1株の染色体DN Aを鋳型に、上記調製したプライマー (例、配列番号4 と5)を用いてPCR反応を行い、DNA断片を増幅さ せる。増幅された断片のDNA配列(例、配列番号6) を決定し、当初設定したアミノ酸配列をコードしている ことを確認後、該断片をプローブとし、染色体DNAの 制限酵素切断産物に対し、サザンハイブリダイゼーショ ンを実施する。目的とするDNAポリメラーゼ遺伝子を 含む断片のおおよその大きさを約4~7Kbpに限定す ることが好ましい。

【0014】更に、約4~7KbpのDNA断片をゲル から回収し、これを用いて、大腸菌にてDNAライブラ リーを作製し、上記記載のPCR増幅DNA断片 (例、 配列番号6)をプロープにコロニーハイブリダイゼーシ ョンを行い、クローン株を取得する。

*DNAポリメラーゼをコードする単離されたDNAをベ クターに挿入したDNA組換え発現ペクターを用いて形 質転換された組換え宿主細胞を培養し、(a) 該組換え 宿主細胞を集めた後、破砕し、細胞抽出物を調製し、

(b) 組換え宿主細胞由来の不純蛋白質を除去する工程 を含むことを特徴とする超好熱始原菌KOD1由来DN Aポリメラーゼを精製する方法である。

【0010】本発明において使用する超好熱始原菌の1 種であるKOD1は、鹿児島県小宝島の硫気抗から単離 した菌株である。該菌株の菌学的性質を以下に記載す

球菌・二連球菌、鞭毛あり

65~100℃

95℃

5~9

2~3%

従属栄養

嫌気性

エーテル型 38%

のDNAポリメラーゼ遺伝子は5010塩基(推定アミ ノ酸1670個)から構成されている(配列番号1)。 他のDNAポリメラーゼと比較したところ、本発明の遺 伝子には真核生物型であるαDNAポリメラーゼの保存 領域、Region1~5が存在している。また該遺伝 子のN末端側に3'→5'エキソヌクレア-ゼモチ-フ であるEXO1, 2, 3が存在している。超好熱始原菌 KOD1株由来の耐熱性DNAポリメラーゼ遺伝子の保 存領域、Region1,2内には、各々介在配列が存 在しており、かつオープンリーディングフレーム(OR F)の保存された形でつながっている。

【0016】超好熱始原菌KOD1株の耐熱性DNAポ リメラーゼ遺伝子を、既知酵素であるピロコッカス・フ リオサス(Pyrococcus furiosus) 由来のPfuDNAポ リメラーゼ遺伝子(特開平 5-328969 号公報)、及びサ ーモコッカス・リトラリス(Thermococcus litoralis)由 来のTli (Vent) DNAポリメラーゼ遺伝子(特 開平 6-7160 号公報) と比較すると、本発明のKOD1 株の遺伝子には介在配列が存在するが、上記PfuDN Aポリメラーゼの遺伝子には介在配列は存在せず、また TliDNAポリメラーゼ遺伝子には、2種の介在配列 が存在するものの、その存在箇所は各々保存領域である Region 2, 3の内であり、本発明のKOD 1株の 耐熱性DNAポリメラーゼ遺伝子内の介在配列の存在箇 所とは大きく異なっている(図4参照)。

【0017】本発明の遺伝子は超好熱始原菌KOD1由 来のDNAポリメラーゼをコードするDNAである。該 【0015】本発明においてクローン化したKOD1株 50 DNAの一例は配列番号1または2に記載されるアミノ

....

酸配列をコードする塩基配列を含有する。また、このよ うなDNAは配列番号1または3に記載される塩基配列 またはその一部分を含有する。本発明の超好熱始原菌K OD1株由来の耐熱性DNAポリメラーゼを大陽菌で発 現させるため、配列番号1に示される塩基配列の137 4~2453bp、2708~4316bpの介在配列 をPCR遺伝子融合法により取り除き、完全な形のDN Aポリメラーゼ遺伝子を構築する。 具体的には、介在配 列を含むクローン化した遺伝子を3組のプライマーの組 み合わせによりPCR反応を行い、介在配列により分断 10 される3断片を増幅する。ここで使用するプライマーを 設計する際、その末端に結合すべき断片の一部をその 5、端に含ませておく。次いで、結合すべき断片同志を 用いてその末端の重複する配列を利用してPCR反応を 行い、各々断片を結合する。更に得られた2種の断片を 用い同様にPCR反応を行い、介在配列を含まないKO D1株由来のDNAポリメラーゼ遺伝子を含まない、完 全な形のDNAポリメラーゼ遺伝子を得る。

【0018】本発明において使用するベクターは、KO D1由来の耐熱性DNAポリメラーゼのクローニングお 20 よび発現を可能とするものであれば、いかなるものでも よく、例えばファージおよびプラスミドが挙げられる。 プラスミドとしては、T7プロモーターで誘導発現が可 能なプラスミドベクター、例えばpET-8cなどを挙 げることができる。また別なプラスミドの例としては、 pUC19, pBR322, pBluescript, pSP73、pGW7、pET3A、pET11Cなど がある。ファージとしては、たとえばλgt11、λD ASH、AZapIIなどが挙げられる。本発明におい て使用する宿主細胞としては、大腸菌、酵母などが挙げ られる。大腸菌としては、例えばJM109、101、 XL1、PR1、BL21 (DE3) plysSなどが 挙げられる。本発明では上記KOD1由来の耐熱性DN Aポリメラーゼをコードする遺伝子を上記ベクターに挿 入して組換え発現ペクターとし、更に、この組換え発現 ベクターにて宿主細胞を形質転換する。

【0019】本発明の製造法では、上記組換え宿主細胞 を培養して、KOD1株由来の耐熱性DNAポリメラー ゼ遺伝子を誘導発現させる。 組換え宿主細胞の培養に使 用する培地ならびに条件は常法に従う。具体例として は、KOD1株由来の介在配列を含まない完全な形のD NAポリメラーゼ遺伝子を含むpET-8cプラスミド により形質転換された大腸菌を、例えばTB培地にて培 養し、誘導処理する。T7プロモーターの誘導処理はイ ソプロピオチ- β-D- ガラクトシドの添加により行なう ことが好ましい。

【0020】本発明の精製法では、組換え宿主細胞を培 養した後、(a)組換え宿主細胞を集めた後、破砕し、 細胞抽出物を調製し、(b)宿主細胞由来の不純蛋白質

耐熱性DNAポリメラーゼは、宿主菌体を培地で培養・ 誘導処理後、培養液から遠心分離等にて分離・回収す る。該菌体を緩衝液に再懸濁した後、超音波処理、ダイ ノミル・フレンチプレス等により菌体を破砕する。次い で、熱処理を実施し、上清より耐熱性DNAポリメラー ゼを回収する。菌体破砕方法は、超音波処理、ダイノミ ル・フレンチプレス法などが好ましい。宿主細胞由来の 不純タンパク質を除去する工程の1つとして、熱処理が 好ましい。熱処理条件は70℃以上、好ましくは90℃ 以上である。他の不純タンパク質の除去法としては各種 クロマトグラフィーなどを実施する。

【0021】この様にして取得した超好熱始原菌KOD 1株由来の耐熱性DNAポリメラーゼの分子量は、約9 0KDaである(図2参照)。

【0022】また、この耐熱性DNAポリメラーゼを用 いポリメラーゼ連鎖反応を実施すると、十分な目的DN A断片の増幅が確認される(図3参照)。

[0023]

【発明の効果】本発明により取得される超好熱始原菌由 来のDNAポリメラーゼは、高い熱安定性を有し、ポリ メラーゼ連鎖反応等に適した酵素である。

[0024]

【実施例】次に本発明を実施例を用いて説明する。 実施例1

超好熱始原菌KOD1株由来DNAポリメラーゼ遺伝子 のクローニング

鹿児島県小宝島にて単離した超好熱始原菌KOD1株を 95℃にて培養後、菌体を回収した。得られた菌体から 常法に従い超好熱始原菌KOD1株の染色体DNAを調 製した。Pyrococcus furiosus 由来のDNAポリメラー ゼ(Pfuポリメラーゼ)の保存領域アミノ酸配列に基 づき、2種のプライマー(5'-GGATTAGTATAGTGCCAATGGAA GGCGAC-3'(配列番号 4), 5'-GAGGGCGAAGTTTATTCCGAGCTT -3'(配列番号5) を合成した。この2種のプライマーを 使用し、調製した染色体DNAを鋳型として、PCR反 広を行った。

【0025】PCR増幅DNA断片の塩基配列(配列番 号6)を決定し、アミノ酸配列(配列番号7)を決定し た後、この増幅DNA断片をプロープとして、KOD1 株染色体DNA制限酵素処理産物に対してサザンハイブ リダイゼーションを行い、DNAポリメラーゼをコード する断片のサイズを求めた(約1~7 K b p)。さら に、この大きさのDNA断片をアガロースゲルから回収 し、プラスミドpBS (ストラタジーン社製) に挿入 し、これらの混合物により大腸菌(B.coli JM109)を形質 転換して、ライブラリーを作製した。サザンハイブリダ イゼーションに使用したプローブ(配列番号6)を用い て、コロニーハイプリダイゼーションを行い、上記ライ プラリーから、KOD1株由来のDNAポリメラーゼ遺 を除去する工程を含む。組換え宿主細胞より産出された 50 伝子を含有すると考えられるクローン株(E.coli JM109/

pBSKOD1) を取得した。

【0026】 実施例2

クローン断片の塩基配列の決定

実施例1で取得したクローン株、E. coli JM109/pBSK0D1 よりプラスミド、BSK0D1を回収し、常法に従い塩基配列 (配列番号1)を決定した。さらに求められた塩基配列 からアミノ酸配列を推定した。KOD1株由来のDNA ポリメラーゼ遺伝子は5010塩基からなり、1670 個のアミノ酸がコードされていた。

7

【0027】実施例3

組換え発現ベクターの構築

完全なポリメラーゼ遺伝子を作成するため、2箇所の介 在配列部分(1374~2453bp、2708~43 16bp)をPCR融合法により取り除いた。PCR融 合法では、クローン株より回収したプラスミドを鋳型 に、3組のプライマー(配列番号8~13)を組み合わ せて、各々PCRを行い、介在配列を除いた3断片を増 幅した。この際、PCRに用いるプライマーは、他の断 片と結合する側に結合相手と同様な配列がくるように設 計した。また、両端には別々の制限酵素サイト(N末端 20 側:EcoRV、C末端側:BamHI) が創出される ように設計した。次いで、PCR増幅断片中、構造上中 央に位置する断片と、N末端側に位置する断片を混合 し、PCRを各々の断片をプライマーとして行った。ま た、同様に構造上、中央に位置する断片と、C末端側に 位置する断片を混合し、PCRを各々の断片をプライマ ーとして行った。このようにして得られた2種の断片を 用いて再度PCRを行い、介在配列が取り除かれ、N末 端にEcoRV、C末端にBamHIサイトを有するK OD1株由来のDNAポリメラーゼをコードする完全な 30 形の遺伝子断片を取得した。更に、同遺伝子をT7プロ モーターで誘導可能な発現ペクター、pET-8cのN col/BamHlサイト、先に創出した制限酵素サイ トを利用し、サブクローニングして、組換え発現ベクタ - (pET-pol) を得た。

[0028] 実施例4

KOD1由来DNAポリメラーゼの発現と精製

実施例3で取得した組換え発現ベクター(pET-pol)を用いて大腸菌(E.coli JM109)を形質転換し、得られた形質転換体をTB培地(Molecular Cloning, p.A.2, 1989に記載)で培養し、集菌1時間前にT7プロモーターの誘導処理をイソプロピオチーβ-D-ガラクトシドの添加により行った。培養液より菌体を違心分離により回収した。緩衝液に再懸濁した後、超音波処理によって菌体を破砕し、細胞抽出物を得た。さらに宿主細胞由来の不純タンパク質を除去するために、細胞破砕液を94℃にて20分間処理し、宿主細胞由来の不純タンパク質を不溶化した。不溶画分を遠心分離して除去し、KOD1株由来の耐熱性DNAポリメラーゼを得た。

【0029】実施例5

KOD 1由来耐熱性DNAポリメラーゼの精製

実施例4で得られたKOD1由来耐熱性DNAポリメラーゼの分子量をSDS-PAGE法によって求めたところ、約86~92kDaであった(図2)。また、実施例4で得たKOD1由来の耐熱性DNAポリメラーゼと既知の鋳型・プライマーを用いてPCRを実施したところ、サーモコッカス・リトラリス(Thermococcus litora lis)由来の耐熱性DNAポリメラーゼを用いた場合と同様に標的とするDNA断片が確認され(図3)、高い20 熱安定性DNAポリメラーゼ活性が確認された。

【0030】比較例1

本発明の超好熱姶原菌KOD1と類縁菌であると思われ るピロコッカス・フリオサス(Pyrococcus furiosus) ま たはサーモコッカス・リトラリス(Thermococcus litora lis) 由来の耐熱性DNAポリメラーゼ遺伝子との比較 本発明の超好熱始原南KOD1由来のDNAポリメラー ゼ遺伝子(配列番号3)、ピロコッカス・フリオサス(P yrococcus furiosus) 由来の耐熱性DNAポリメラーゼ 遺伝子(特開平 5-328969 号公報)、サーモコッカス・ リトラリス (Thermococcus litoralis) 由来の耐熱性D NAポリメラーゼ遺伝子 (特開平 6-7160 号公報) のD NA配列からアミノ酸配列を推定し、比較検討した。本 発明のKOD1由来のDNAポリメラーゼは、真核生物 型である α DNA ポリメラーゼの保存領域である Reg ion1~5が存在していた。またN末端側には3'→ 5'エキソヌクレアーゼモチーフであるEXO1, 2, 3が存在していた。しかし、αDNAポリメラーゼ保存 領域Region1とRegion2の内には、各々介 在配列IVS-A、IVS-Bが存在していた(図4参 照)。一方、ピロコッカス・フリオサス(Pyrococcus fu riosus) 由来の耐熱性DNAポリメラーゼであるPfu ポリメラーゼには介在配列が存在しなかった。またサー モコッカス・リトラリス (Thermococcus litoralis) 由 来の耐熱性DNAポリメラーゼであるVentポリメラ ーゼでは、αDNAポリメラーゼ保存領域Region 2とRegion3の内に、介在配列IVS1とIVS 2が認められた (図4参照)。

[0031]

【配列表】

40 配列番号1

配列の長さ:5342

配列の型:核酸 (DNA)

鎖の数:2本鎖

トロポジー:直鎖状

配列の種類:c DNA

起源:超好熱始原菌

株名:KOD1

配列の特徴

156-5165 P CDS

50 1374-2453 介在配列

--507--

2708-4316 介在配列

교

GCTTGAGGGC CTGCGGTTAT GGGACGTTGC AGTTTGCGCC TACTCAAAGA TGCCGGTTTT ATAACGGAGA AAAATGGGGA GCTATTACGA TCTCTCCTTG ATGTGGGGTT TACAATAAAG 120 CCTGGATTGT TCTACAAGAT TATGGGGGAT GAAAG ATG ATC CTC GAC ACT GAC Met Ile Leu Asp Thr Asp 1 TAC ATA ACC GAG GAT GGA AAG CCT GTC ATA AGA ATT TTC AAG AAG GAA 221 Tyr Ile Thr Glu Asp Gly Lys Pro Vai Ile Arg Ile Phe Lys Lys Glu 10 15 AAC GGC GAG TTT AAG ATT GAG TAC GAC CGG ACT TTT GAA CCC TAC TTC 269 Asn Gly Glu Phe Lys Ile Glu Tyr Asp Arg Thr Phe Glu Pro Tyr Phe 30 TAC GCC CTC CTG AAG GAC GAT TCT GCC ATT GAG GAA GTC AAG AAG ATA 317 Tyr Ala Leu Leu Lys Asp Asp Ser Ala Ile Glu Glu Val Lys Lys Ile 45 50 ACC GCC GAG AGG CAC GGG ACG GTT GTA ACG GTT AAG CGG GTT GAA AAG 365 Thr Ala Glu Arg His Gly Thr Val Val Thr Val Lys Arg Val Glu Lys 60 65 GTT CAG AAG AAG TTC CTC GGG AGA CCA GTT GAG GTC TGG AAA CTC TAC Val Gln Lys Lys Phe Leu Gly Arg Pro Val Glu Val Trp Lys Leu Tyr TTT ACT CAT CCG CAG GAC GTC CCA GCG ATA AGG GAC AAG ATA CGA GAG Phe Thr His Pro Gln Asp Val Pro Ala Ile Arg Asp Lys Ile Arg Glu 90 95 CAT GGA GCA GTT ATT GAC ATC TAC GAG TAC GAC ATA CCC TTC GCC AAG 509 His Gly Ala Val Ile Asp Ile Tyr Glu Tyr Asp Ile Pro Phe Ala Lys 110 CGC TAC CTC ATA GAC AAG GGA TTA GTG CCA ATG GAA GGC GAC GAG GAG 557 Arg Tyr Leu Ile Asp Lys Gly Leu Val Pro Met Glu Gly Asp Glu Glu 125 130 CTG AAA ATG CTC GCC TTC GAC ATT CAA ACT CTC TAC CAT GAG GGC GAG Leu Lys Met Leu Ala Phe Asp Ile Gin Thr Leu Tyr His Glu Gly Glu 145 GAG TTC GCC GAG GGG CCA ATC CTT ATG ATA AGC TAC GCC GAC GAG GAA Glu Phe Ala Glu Gly Pro Ile Leu Met Ile Ser Tyr Ala Asp Glu Glu GGG GCC AGG GTG ATA ACT TGG AAG AAC GTG GAT CTC CCC TAC GTT GAC 701 Gly Ala Arg Val Ile Thr Trp Lys Asn Val Asp Leu Pro Tyr Val Asp 170 175 GTC GTC TCG ACG GAG AGG GAG ATG ATA AAG CGC TTC CTC CGT GTT GTG 749 Val Val Ser Thr Glu Arg Glu Met Ile Lys Arg Phe Leu Arg Val Val 190 AAG GAG AAA GAC CCG GAC GTT CTC ATA ACC TAC AAC GGC GAC AAC TTC 797 Lys Glu Lys Asp Pro Asp Val Leu Ile Thr Tyr Asm Gly Asp Asm Phe 205 210 GAC TTC GCC TAT CTG AAA AAG CGC TGT GAA AAG CTC GGA ATA AAC TTC 845 Asp Phe Ala Tyr Leu Lys Lys Arg Cys Glu Lys Leu Gly Ile Asn Phe 215 220 225 GCC CTC GGA AGG GAT GGA AGC GAG CCG AAG ATT CAG AGG ATG GGC GAC

Ala Leu Gly Arg Asp Gly Ser Glu Pro Lys Ile Gln Arg Met Gly Asp 245 AGG TTT GCC GTC GAA GTG AAG GGA CGG ATA CAC TTC GAT CCT TAY CCT 250 250 CTG ATA AGA CGG ACG ATA AAC CTG CCC ACA TAC ACG CTT GAG GCC GTT 989 Val Ile Arg Arg Thr Ile Asn Leu Pro Thr Tyr Thr Leu Glu Ala Val 265 TAT GAA GCC GTC TTC GGT CAG CCG AAG ACC ATA CAC TTC GAG GCC GTT 989 Val Ile Arg Arg Thr Ile Asn Leu Pro Thr Tyr Thr Leu Glu Ala Val 265 270 TAT GAA GCC GTC TTC GGT CAG CCG AAG GAG GTT TAC GCT GAG GCC GTT 989 ATA ACA CCA GCC TG GG CAA ACC GCC GAG AAC GTT GAG AGA GTC GCC CCC 1085 Thr Gaa GCC GTC TTC GGT CAG CGC GAG AAC CTT GAG AGA GTC GCC CCC 1085 Thr Gaa GCC GTC TTC GGT CAG CGC GAG AAC CTT GAG AGA GTC GCC CCC 1085 Thr Gaa GCC GTC TTC GGT GAG GCC GAG ACC GTC GGC AAG GTC GCC CCC 1085 Thr Gaa GCC GCC GGG GAA GCC GCC GAG ACC GCC GAG ACC GTC GCG AAG AGA GTC GCC CCC 1085 Thr Gaa GAA GAT GAG GTC GCC GAG ACC GCC GAG ACC GTC GCG AAG GAG TTC 1133 Thr Ser Met Glu Asp Ala Lys Val Thr Tyr Glu Leu Gly Lys Glu Phe 315 TAC TCG ATG GAG GCC CAG CTT TCT CCC TTA ATC GCC CAG TCC CTC TGG 1181 Leu Pro Met Glu Ala Glu Leu Ser Arg Leu Ile Gly Gln Ser Leu Trp 330 TAS ACC ACC ACC ACC ACC ACC ACC TCC ACC ACC
Arg Phe Ala Val Glu Val Lys Gly Arg Ile His Phe Asp Leu Tyr Pro 250 255 260 260 GTG ATA AGA CGG ACG ATA AAC CTG CCC ACA TAC ACG CTT GAG GCC GTT 989 Val Ile Arg Arg Thr Ile Asn Leu Pro Thr Tyr Thr Leu Glu Ala Val 265 270 275 TAT GAA GCC GTC TTC GGT CAG CCG AAG GAG AAG GTT TAC GCT GAG GAA 1037 Tyr Glu Ala Val Phe Gly Gln Pro Lys Glu Lys Val Tyr Ala Glu Glu 285 290 ATA ACA CCA GCC TGG GAA ACC GGC GAG AAC CTT GAG AGA GTC GCC CGC 1085 The Thr Pro Ala Trp Glu Thr Gly Glu Asn Leu Glu Arg Val Ala Arg 295 300 305 310 TAC TGG ATG GAA GAT GCG AAG GTC ACA TAC GAG CTT GGG AAG GTC GCC CGC 1133 Tyr Ser Met Glu Asp Ala Lys Val Thr Tyr Glu Leu Gly Lys Glu Phe 315 320 325 CTT CCG ATG GAG GCC CAG CTT TCT CGC TTA ATC GGC CAG TCC CTC TGG 1181 Leu Pro Met Glu Ala Gln Leu Ser Arg Leu Ile Gly Gln Ser Leu Trp 330 335 340 GAC GTC TCC CGC TCC AGC ACT GGC AAC CTC GTT GAG TGC TTC CTC TCC ASp Val Ser Arg Ser Ser Thr Gly Asn Leu Val Glu Trp Phe Leu Leu 345 350 355 AGG AAG GCC TAT GAG AGA AGA GGC CAG AAC CTC GTT GAG GAC CCC GAT GAA ARG Lys Ala Tyr Glu Arg Asn Glu Leu Ala Pro Asn Lys Pro Asp Glu 366 365 370 AAG GAG CTG GCC AGA AGA GAG GGC CAG AGA GAG GCC TAT GTA AAA Lys Glu Leu Ala Arg Arg Arg Grg Glo Leu Ala Pro Asn Lys Pro Asp Glu 367 380 385 390 GAG CCC GAG AGA AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA AGG CCC GAG AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA AGG CCC GAG AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA AGG CCC GAG AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA AGG CCC GAG AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA AGG CCC GAG AGA AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA AGG CCC GAG CAG AGA GAG GTT GTG GG GAG AAC ATA GTG TAC CTA GAT TTT AGA AGG CCC GAG AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA AGG CCC GAG CAG AGA GAG GTT GTG GGG GAG AAC ATA GTG TAC CTA GAT TTT AGA AGG CCC GAG CAG AGA GAG GTT GTG GAG AAC ATA GTG TAC CTA GAT TTT AGA AGG CCC GAG CAG AGA GAG GTT GTG GAG AAC ATA GTG TAC CTA GAT TTT AGA AGG CCC GAG CAG AGA GAG GTT GTG GAG AAC ATA GTG
250 255 260 285 260 285 260 285 260 285 260 265 270 275
Val Lie Arg Arg Thr Lie Asn Leu Pro Thr Tyr The Leu Giu Ala Val 265 270 275 275 TAT GAA GCC GTC CCG CCG CCG AAG GTT TAC GCT GAA AA 1037 Tyr Giu Ala Val Phe Giu Giu Fro Lys Val Tyr Ala GIu Fro GIu Fro GIu Ala Arg GIu Fro Ala Arg GCC CGC GGC AAG GTC AAC ATC GGG AAG GTC CTC GGG AAG GTC ATC GGG AAG GTC CTC GGG AAG GTC CTC TCC GTC
265
TAT GAA GCC GTC TTC GGT CAG CCG AAG GAG GAG AAG GTT TAC GCT GAG GAA 1037 Tyr Glu Ala Val Phe Gly Gln Pro Lys Glu Lys Val Tyr Ala Glu Glu 280 285 290 ATA ACA CCA GCC TGG GAA ACC GGC GAG AAC CTT GAG AGA GTC GCC CGC 1085 Ile Thr Pro Ala Trp Glu Thr Gly Glu Asn Leu Glu Arg Val Ala Arg 295 300 305 310 TAC TCG ATG GAA GAT GCG AAG GTC ACA TAC GAG CTT GGC AAG GAG TTC 1133 Tyr Ser Met Glu Asp Ala Lys Val Thr Tyr Glu Leu Gly Lys Glu Phe 315 320 325 CTT CCG ATG GAG GCC CAG CTT TCT CGC TTA ATC GGC CAG TCC CTC TGG 1181 Leu Pro Met Glu Ala Gln Leu Ser Arg Leu Ile Gly Gln Ser Leu Trp 330 335 340 GAC GTC TCC CGC TCC AGC ACT GGC AAC CTC GTT GAG TGG TTC CTC CTC Asp Val Ser Arg Ser Ser Thr Gly Asn Leu Val Glu Trp Phe Leu Leu 345 350 366 367 AAG GAG GCC TAT GAG AGA GAG GAG GAG GAG GAG GAG AGA GAG
Tyr Glu Ala Val Phe Gly Gln Pro Lys Glu Lys Val Tyr Ala Glu Glu 280
280
ATA ACA CCA GCC TGG GAA ACC GGC GAG AAC CTT GAG AGA GTC GCC CGC 1085 Ile Thr Pro Ala Trp Glu Thr Gly Glu Asn Leu Glu Arg Val Ala Arg 295 300 305 310 TAC TCG ATG GAA GAT GCG AAG GTC ACA TAC GAG CTT GGG AAG GAG TTC 1133 Tyr Ser Met Glu Asn Ala Lys Val Thr Tyr Glu Leu Gly Lys Glu Phe 315 320 325 CTT CCG ATG GAG GCC CAG CTT TCT CGC TTA ATC GCC CAG TCC CTC TGG 1181 Leu Pro Met Glu Ala Gln Leu Ser Arg Leu Ile Gly Gln Ser Leu Trp 330 335 340 GAC GTC TCC CGC TCC AGC ACT GGC AAC CTC GTT GAG TGG TTC CTC CTC TGG 1229 Asn Val Ser Arg Ser Ser Thr Gly Asn Leu Val Glu Trp Phe Leu Leu 345 350 355 AGG AAG GCC TAT GAG AGG AAT GAG CTG CCC CCG AAC AAG CCC GAT GAA 1277 Arg Lys Ala Tyr Glu Arg Asn Glu Leu Ala Pro Asn Lys Pro Asp Glu 360 365 370 AAG GAG CTG GCC AGA AGA CGG CAG AGC TAT GAA AGA CCC GAG GAG CCC GAG GAG GAG GAG
The Pro Ala Trp Glu Thr Gly Glu Asn Leu Glu Arg Val Ala Arg
300 305 310
TAC TCG ATG GAA GAT GCG AAG GTC ACA TAC GAG CTT GGG AAG GAG TTC Tyr Ser Met Glu Asp Ala Lys Val Thr Tyr Glu Leu Gly Lys Glu Phe 315 320 325 CTT CCG ATG GAG GCC CAG CTT TCT CGC TTA ATC GGC CAG TCC CTC TGG 1181 Leu Pro Met Glu Ala Gln Leu Ser Arg Leu Ile Gly Gln Ser Leu Trp 330 335 340 GAC GTC TCC CGC TCC AGC ACT GGC AAC CTC GTT GAG TGG TTC CTC CTC 1229 Asp Val Ser Arg Ser Ser Thr Gly Asn Leu Val Glu Trp Phe Leu Leu 345 350 355 AGG AAG GCC TAT GAG AGG AAT GAG CTG GCC CCG AAC AAG CCC GAT GAA 1277 Arg Lys Ala Tyr Glu Arg Asn Glu Leu Ala Pro Asn Lys Pro Asp Glu 360 365 370 AAG GAG CTG GCC AGA AGA CGG CAG AGC TAT GAA GGA GGC TAT GTA AAA 1325 Lys Glu Leu Ala Arg Arg Arg Gln Ser Tyr Glu Gly Gly Tyr Val Lys 375 380 385 390 GAG CCC GAG AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA 1373 Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile Val Tyr Leu Asp Phe Arg 395 400 405 TCC CAT CCA GCC GAT ACG AAG GTT GTC GTC AAG GGG AAG GGG ATT ATA 1421 Cys His Pro Ala Asp Thr Lys Val Val Val Lys Gly Lys Gly Ile Ile 410 415 420
Tyr Ser Net Glu Asp Ala Lys Val Thr Tyr Glu Leu Gly Lys Glu Phe 315
Signature Sign
CTT CCG ATG GAG GCC CAG CTT TCT CGC TTA ATC GGC CAG TCC CTC TGG 1181 Leu Pro Met Glu Ala Gln Leu Ser Arg Leu Ile Gly Gln Ser Leu Trp 330 340 GAC GTC TCC CGC TCC AGC ACT GGC AAC CTC GTT GAG TGG TTC CTC CTC 1229 Asp Val Ser Arg Ser Ser Thr Gly Asn Leu Val Glu Trp Phe Leu Leu 345 350 355 AGG AAG GCC TAT GAG AGG AAT GAG CTG GCC CCG AAC AAG CCC GAT GAA 1277 Arg Lys Ala Tyr Glu Arg Asn Glu Leu Ala Pro Asn Lys Pro Asp Glu 360 365 370 AAG GAG CTG GCC AGA AGA CGG CAG AGC TAT GAA GGA GGC TAT GTA AAA 1325 Lys Glu Leu Ala Arg Arg Arg Gln Ser Tyr Glu Gly Gly Tyr Val Lys 375 380 385 390 GAG CCC GAG AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA 1373 Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile Val Tyr Leu Asp Phe Arg 395 TGC CAT CCA GCC GAT ACG AAG GTT GTC GTC AAG GGG AAG GGG ATT ATA 1421 Cys His Pro Ala Asp Thr Lys Val Val Val Lys Gly Lys Gly Ile Ile 410 415 420
Leu Pro Net Glu Ala Gln Leu Ser Arg Leu Ile Gly Gln Ser Leu Trp 330 335 340 GAC GTC TCC CGC TCC AGC ACT GGC AAC CTC GTT GAG TGG TTC CTC CTC 1229 Asp Val Ser Arg Ser Ser Thr Gly Asn Leu Val Glu Trp Phe Leu Leu 345 350 355 AGG AAG GCC TAT GAG AGG AAT GAG CTG GCC CCG AAC AAG CCC GAT GAA 1277 Arg Lys Ala Tyr Glu Arg Asn Glu Leu Ala Pro Asn Lys Pro Asp Glu 360 365 370 AAG GAG GAG CTG GCC AGA AGA AGA CGG CAG AGC TAT GAA GGA GGC TAT GTA AAA 1325 380 385 390 GAG CCC GAG AGA GAG GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA 1373 390 385 390 GAG CCC GAG AGA GAG GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA 1373 395 400 405 TGC CAT CCA GCC GAT ACG AAG GTT GTC GTC AGG GTC AGG GGG ATT ATA 1421 420 420
330 335 340 GAC GTC TCC CGC TCC AGC ACT GGC AAC CTC GTT GAG TGG TTC CTC CTC 1229 Asp Val Ser Arg Ser Ser Thr Gly Asn Leu Val Glu Trp Phe Leu Leu 345 350 355 AGG AAG GCC TAT GAG AGG AAT GAG CTG GCC CCG AAC AAG CCC GAT GAA 1277 Arg Lys Ala Tyr Glu Arg Asn Glu Leu Ala Pro Asn Lys Pro Asp Glu 360 365 370 AAG GAG CTG GCC AGA AGA AGA CGG CAG AGC TAT GAA GGA GGC TAT GTA AAA 1325 Lys Glu Leu Ala Arg Arg Arg Gln Ser Tyr Glu Gly Gly Tyr Val Lys 375 380 385 390 GAG CCC GAG AGA GAG GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA 1373 Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile Val Tyr Leu Asp Phe Arg 395 TGC CAT CCA GCC GAT ACG AAG GTT GTC GTC AAG GGG AAG GGG ATT ATA 1421 Cys His Pro Ala Asp Thr Lys Val Val Val Lys Gly Lys Gly Ile Ile 410 415 420
AAG GTC TCC CGC TCC AGC ACT GGC AAC CTC GTT GAG TGG TTC CTC CTC 1229 ASP Val Ser Arg Ser Ser Thr Gly Asn Leu Val Glu Trp.Phe Leu Leu 345 350 355 AGG AAG GCC TAT GAG AGG AAT GAG CTG GCC CCG AAC AAG CCC GAT GAA 1277 1277 Arg Lys Ala Tyr Glu Arg Asn Glu Leu Ala Pro Asn Lys Pro Asp Glu 360 365 370 AAG GAG CTG GCC AGA AGA AGA CGG CAG AGC TAT GAA GGA GGC TAT GTA AAA 1325 1325 Lys Glu Leu Ala Arg Arg Arg Gln Ser Tyr Glu Gly Gly Tyr Val Lys 375 380 385 390 GAG CCC GAG AGA GAG GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA 1373 Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile Val Tyr Leu Asp Phe Arg 395 400 405 TGC CAT CCA CCC GAT ACG AAG GTT GTC GTC AAG GGG AAG GGG ATT ATA 1421 415 420
ASP Val Ser Arg Ser Ser Thr Gly Asn Leu Val Glu Trp Phe Leu Leu 345 AGG AAG GCC TAT GAG AGG AAT GAG CTG GCC CCG AAC AAG CCC GAT GAA 1277 Arg Lys Ala Tyr Glu Arg Asn Glu Leu Ala Pro Asn Lys Pro Asp Glu 360 365 AAG GAG CTG GCC AGA AGA CGG CAG AGC TAT GAA GGA GGC TAI GTA AAA 1325 Lys Glu Leu Ala Arg Arg Arg Gln Ser Tyr Glu Gly Gly Tyr Val Lys 375 380 380 385 390 GAG CCC GAG AGA GAG GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA 1373 Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile Val Tyr Leu Asp Phe Arg 395 TGC CAT CCA GCC GAT ACG AAG GTT GTC GTC AAG GGG AAG GGG ATT ATA 1421 Cys His Pro Ala Asp Thr Lys Val Val Val Lys Gly Lys Gly Ile Ile 410 415 420
345 350 355 AGG AAG GCC TAT GAG AGG AAT GAG CTG GCC CCG AAC AAG CCC GAT GAA 1277 Arg Lys Ala Tyr Glu Arg Asn Glu Leu Ala Pro Asn Lys Pro Asp Glu 360 365 370 AAG GAG CTG GCC AGA AGA CGG CAG AGC TAT GAA GGA GGC TAT GTA AAA 1325 Lys Glu Leu Ala Arg Arg Arg Gln Ser Tyr Glu Gly Gly Tyr Val Lys 375 380 385 390 GAG CCC GAG AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA 1373 Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile Val Tyr Leu Asp Phe Arg 395 400 405 TCC CAT CCA GCC GAT ACG AAG GTT GTC GTC AAG GGG AAG GGG ATT ATA 1421 Cys His Pro Ala Asp Thr Lys Val Val Val Lys Gly Lys Gly Ile Ile 410 415 420
AGG AAG GCC TAT GAG AGG AAT GAG CTG GCC CCG AAC AAG CCC GAT GAA 1277 Arg Lys Ala Tyr Glu Arg Asn Glu Leu Ala Pro Asn Lys Pro Asp Glu 360 365 370 AAG GAG CTG GCC AGA AGA CGG CAG AGC TAT GAA GGA GGC TAT GTA AAA 1325 Lys Glu Leu Ala Arg Arg Arg Gln Ser Tyr Glu Gly Gly Tyr Val Lys 375 380 385 390 GAG CCC GAG AGA GAG GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile Val Tyr Leu Asp Phe Arg 395 400 TCC CAT CCA GCC GAT ACG AAG GTT GTC GTC AAG GGG AAG GGG ATT ATA 1421 Cys His Pro Ala Asp Thr Lys Val Val Val Lys Gly Lys Gly Ile Ile 410 415 420
Arg Lys Ala Tyr Glu Arg Asn Glu Leu Ala Pro Asn Lys Pro Asp Glu 360 365 370 AAG GAG CTG GCC AGA AGA CGG CAG AGC TAT GAA GGA GGC TAT GTA AAA 1325 Lys Glu Leu Ala Arg Arg Arg Gln Ser Tyr Glu Gly Gly Tyr Val Lys 375 380 385 390 GAG CCC GAG AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA 1373 Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile Val Tyr Leu Asp Phe Arg 395 400 405 TCC CAT CCA GCC GAT ACG AAG GTT GTC GTC AAG GGG AAG GGG ATT ATA 1421 Cys His Pro Ala Asp Thr Lys Val Val Val Lys Gly Lys Gly Ile Ile 410 415 420
360 365 370 AAG GAG CTG, GCC AGA AGA CGG CAG AGC TAT GAA GGA GGC TAT GTA AAA 1325 Lys Glu Leu Ala Arg Arg Arg Gln Ser Tyr Glu Gly Gly Tyr Val Lys 375 380 385 390 GAG CCC GAG AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA 1373 Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile Val Tyr Leu Asp Phe Arg 395 400 405 TGC CAT CCA GCC GAT ACG AAG GTT GTC GTC AAG GGG AAG GGG ATT ATA 1421 Cys His Pro Ala Asp Thr Lys Val Val Val Lys Gly Lys Gly Ile Ile 410 415 420
AAG GAG CTG, GCC AGA AGA CGG CAG AGC TAT GAA GGA GGC TAT GTA AAA 1325 Lys Glu Leu Ala Arg Arg Arg Gln Ser Tyr Glu Gly Gly Tyr Val Lys 375 380 385 390 GAG CCC GAG AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA 1373 Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile Val Tyr Leu Asp Phe Arg 395 400 405 TGC CAT CCA GCC GAT ACG AAG GTT GTC GTC AAG GGG AAG GGG ATT ATA 1421 Cys His Pro Ala Asp Thr Lys Val Val Val Lys Gly Lys Gly Ile Ile 410 415 420
Lys Glu Leu Ala Arg Arg Arg Gln Ser Tyr Glu Gly Gly Tyr Val Lys 375 380 385 390 GAG CCC GAG AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAT TTT AGA 1373 Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile Val Tyr Leu Asp Phe Arg 395 400 405 TGC CAT CCA GCC GAT ACG AAG GTT GTC GTC AAG GGG AAG GGG ATT ATA CYS His Pro Ala Asp Thr Lys Val Val Val Lys Gly Lys Gly Ile Ile 410 415 420
375 380 385 390 GAG CCC GAG AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAI TTT AGA 1373 Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile Val Tyr Leu Asp Phe Arg 395 400 405 TGC CAT CCA GCC GAT ACG AAG GTT GTC GTC AAG GGG AAG GGG ATT ATA 1421 Cys His Pro Ala Asp Thr Lys Val Val Val Lys Gly Lys Gly Ile Ile 410 415 420
GAG CCC GAG AGA GGG TTG TGG GAG AAC ATA GTG TAC CTA GAI TTT AGA 1373 Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile Val Tyr Leu Asp Phe Arg 395 400 405 TGC CAT CCA GCC GAT ACG AAG GTT GTC GTC AAG GGG AAG GGG ATT ATA Cys His Pro Ala Asp Thr Lys Val Val Val Lys Gly Lys Gly Ile Ile 410 415 420
Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile Val Tyr Leu Asp Phe Arg 395 400 405 TGC CAT CCA GCC GAT ACG AAG GTT GTC GTC AAG GGG AAG GGG ATT ATA 1421 Cys His Pro Ala Asp Thr Lys Val Val Val Lys Gly Lys Gly Ile Ile 410 415 420
TGC CAT CCA GCC GAT ACG AAG GTT GTC GTC AAG GGG AAG GGG ATT ATA 1421 Cys His Pro Ala Asp Thr Lys Val Val Val Lys Gly Lys Gly Ile Ile 410 415 420
Cys His Pro Ala Asp Thr Lys Val Val Val Lys Gly Lys Gly Ile Ile 410 415 420
410 415 420
AAC ATC AGC GAG GTT CAG GAA GGT GAC TAT GTC CTT GGG ATT GAC GGC 1469
Asn Ile Ser Glu Val Glu Glu Gly Asp Tyr Val Leu Gly Ile Asp Gly
425 430 435
TGG CAG AGA GTT AGA AAA GTA TGG GAA TAC GAC TAC AAA GGG GAG CTT 1517
Trp Glo Arg Val Arg Lys Val Trp Glu Tyr Asp Tyr Lys Gly Glu Leu
440 445 450
GTA AAC ATA AAC GGG TTA AAG TGT ACG CCC AAT CAT AAG CTT CCC GTT 1565
Val Ash Ile Ash Gly Leu Lys Cys Thr Pro Ash His Lys Leu Pro Val
455 460 465 470
GTT ACA AAG AAC GAA CGA CAA ACG AGA ATA AGA GAC AGT CTT GCT AAG 1613
Val Thr Lys Asn Glu Arg Gln Thr Arg Ile Arg Asp Ser Leu Ala Lys
475 480 485 TCT TTC CTT ACT AAA AAA GTT AAG GGC AAG ATA ATA ACC ACT CCC CTT 1661
Ser Phe Leu Thr Lys Lys Val Lys Gly Lys Ile Ile Thr Thr Pro Leu
490 495 500

1	3														14	
TTC TAT	GAA	ATA	GGC	AGA	GCG	ACA	AGT	GAG	AAT	ATT	CCA	GAA	GAA	GAG	1709	
Phe Tyr															2.00	
	505			0		510					515	4.4	0.4	O I U		
CTT CTC		CCA	CAC	CTC	cer		ATA	CT.	TTC	CCT			100	0780	1050	
GTT CTC															1757	
Val Let		GI¥	Giu	Leu		Gly	He	Leu	Leu	Ala	Glu	Gly	Thr	Leu		
520					525					5 30						
TTG AGG	AAA	GAC	GTT	GAA	TAC	TTT	GAT	TCA	TCC	CGC	AAA	AAA	CGG	AGG	1805	
Leu Arg	Lys .	Азр	Vai	Glu	Туг	Phe	Asp	Ser	Ser	Arg	Lys	Lys	Arg	Arg		
535				540					545					550		
ATT TCA	CAC	CAG	TAT	CGT	GTT	GAG	ATA	ACC	ATT	GGG	AAA	GAC	GAG	GAG	1853	
Ile Sei																
			555	_				560			-30		565	-10		
GAG TT1	AGG	GAT		ATC	ACA	TAC	ΔΤΤ		CAC	ССТ	TTC	TTT		ATT	1001	
Glu Phe															1901	
010 110	, 146		мь	110	Int	131		LUC	GIU	мв	rea		GIY	116		
ACT CC	400	570	T	~~~			575					580				
ACT CCA															1949	
Thr Pro		He	Ser	Glu	Lys	Lys	Gly	Thr	Asn	Ala	Val	Thr	Leu	Lys		
	585					590					595					
GTT GCC	AAG	AAG	AAT	GTT	TAT	CTT	AAA	GTC	AAG	GAA	ATT	ATG	GAC	AAC	1997	
Val Ala	Lys	Lys	Asn	Val	Туг	Leu	Lys	Val	Lys	Glu	Ile	Me t	Asp	Asn		
600)				605					610						
ATA GAG	TCC	CTA	CAT	GCC	CCC	TCG	GTT	CTC	AGG	GGA	TTC	TTC	GAA	GGC	2045	
Ile Glu																
615				620					625	•				630		
GAC GGT	TCA	GTA	AAC		CTT	AGG	ACC	ACT		СТТ	CCA	ACC	CAC		2002	
Asp Gly															2093	
imp or	DCI	141		мв	101	ме	NI E		116	141	nia	101		GIA		
ACA ALC		CAC	635	***	4 75-75		ono	640	m^ 4		-		645			
ACA AAG															2141	
Thr Lys	ASB		Irp	Lys	He	Lys		Val	Ser	Lys	Leu	Leu	Ser	G}n		
		650					655					660				
CTT GG1															2189	
Leu Gly	Ile	Pro	His	Gln	Thr	Tyr	Thr	Tyr	Gln	Tyr	Gln	Glu	Asn	Gly		
	665					670					675					
AAA GAT	CGG	AGC	AGG	TAT	ATA	CTG	GAG	ATA	ACT	GGA	AAG	GAC	GGA	TTG	2237	
Lys Asp	Arg	Ser	Arg	Туг	He	Leu	Glu	He	Thr	Gly	Lys	Asp	Gly	Leu		
680					685					690			-			
ATA CTO	TTC	CAA	ACA	СТС	ATT	GGA	TTC	ATC	AGT	GAA	AGA	AAG	AAC	ССТ	2285	
Ile Let															BBOO	
695				700		٠.,	1	110		4,4	1116	Lyo	ASII			
	AAT	AAC	CCA		ጉ ርተ	CAC	ACC	CAA	705	440	***	TTC		710	2000	
CTG CT1															2333	
Leu Lei	i asi	Lÿ8		116	26L	GID	Arg		Me t	Asn	Asn	Leu		Asn		
			715					720					725			
AAT GG/															2381	
Asn Gly	Phe	Туг	Arg	Leu	Ser	Glu	Phe	Asn	Val	Ser	Thr	Glu	Tyr	Tyr		
		730					735					740				
GAG GGG	AAG	GTC	TAT	GAC	TTA	ACT	CTT	GAA	GGA	ACT	CCC	TAC	TAC	TTT	2429	
Glu Gly																
	745		·	•		750			•	-	755	-	_ • -			
GCC AAT		ATA	TTC	ACC	CAT		ፐርር	СТС	TAC	ccc		ATC	ልተሮ	ልተና	2177	
Ala Ast															4111	
TITE USE	,,	116	TCII	111	1112	USH	OCI	LCU	TAL	LUV	OC I	116	116	116		

15 16 760 ACC CAC AAC GTC TCG CCG GAT ACG CTC AAC AGA GAA GGA TGC AAG GAA 2525 Thr His Asn Val Ser Pro Asp Thr Leu Asn Arg Glu Gly Cys Lys Glu 780 785 TAT GAC GTT GCC CCA CAG GTC GGC CAC CGC TTC TGC AAG GAC TTC CCA 2573 Tyr Asp Val Ala Pro Gln Val Gly His Arg Phe Cys Lys Asp Phe Pro 795 800 GGA TTT ATC CCG AGC CTG CTT GGA GAC CTC CTA GAG GAG AGG CAG AAG 2621 Gly Phe Ile Pro Ser Leu Leu Gly Asp Leu Leu Glu Glu Arg Glu Lys 810 815 ATA AAG AAG AAG ATG AAG GCC ACG ATT GAC CCG ATC GAG AGG AAG CTC Ile Lys Lys Lys Met Lys Ala Thr Ile Asp Pro Ile Glu Arg Lys Leu 830 CTC GAT TAC AGG CAG AGG GCC ATC AAG ATC CTG GCA AAC AGC ATC CTA 2717 Leu Asp Tyr Arg Gin Arg Ala Ile Lys Ile Leu Ala Asn Ser Ile Leu 840 845 850 CCC GAG GAA TGG CTT CCA GTC CTC GAG GAA GGG GAG GTT CAC TTC GTC 2765 Pro Glu Glu Trp Leu Pro Val Leu Glu Glu Glu Glu Val His Phe Val 860 865 AGG ATT GGA GAG CTC ATA GAC CGG ATG ATG GAG GAA AAT GCT GGG AAA 2813 Arg Ile Gly Glu Leu Ile Asp Arg Met Met Glu Glu Asn Ala Gly Lys 875 880 GTA AAG AGA GAG GGC GAG ACG GAA GTG CTT GAG GTC AGT GGG CTT GAA 2861 Val Lys Arg Glu Gly Glu Thr Glu Val Leu Glu Val Ser Gly Leu Glu 890 895 GTC CCG TCC TTT AAC AGG AGA ACT AAC AAG GCC GAG CTC AAG AGA GTA 2909 Val Pro Ser Phe Asn Arg Arg Thr Asn Lys Ala Glu Leu Lys Arg Val 910 AAG GCC CTG ATT AGG CAC GAT TAT TCT GGC AAG GTC TAC ACC ATC AGA 2957 Lys Ala Leu Ile Arg His Asp Tyr Ser Gly Lys Val Tyr Thr Ile Arg 920 925 CTG AAG TCG GGG AGG AGA ATA AAG ATA ACC TCT GGC CAC AGC CTC TTC 3005 Leu Lys Ser Gly Arg Arg Ile Lys Ile Thr Ser Gly His Ser Leu Phe 940 945 TCT GTG AGA AAC GGG GAG CTC GTT GAA GTT ACG GGC GAT GAA CTA AAG 3053 Ser Val Arg Asn Gly Glu Leu Val Glu Val Thr Gly Asp Glu Leu Lys 955 960 CCA GGT GAC CTC GTT GCA GTC CCG CGG AGA TTG GAG CTT CCT GAG AGA Pro Gly Asp Leu Val Ala Val Pro Arg Arg Leu Glu Leu Pro Glu Arg 975 970 AAC CAC GTG CTG AAC CTC GTT GAA CTG CTC CTT GGA ACG CCA GAA GAA Asn His Val Leu Asn Leu Val Glu Leu Leu Leu Gly Thr Pro Glu Glu 990 995 GAA ACT TTG GAC ATC GTC ATG ACG ATC CCA GTC AAG GGT AAG AAG AAC 3197 Glu Thr Leu Asp Ile Val Met Thr Ile Pro Val Lys Gly Lys Lys Asn 1005 1010 TTC TTT AAA GGG ATG CTC AGG ACT TTG CGC TGG ATT TTC GGA GAG GAA 3245 Phe Phe Lys Gly Met Leu Arg Thr Leu Arg Trp Ile Phe Gly Glu Glu 1020 1025 AAG AGG CCC AGA ACC GCG AGA CGC TAT CTC AGG CAC CTT GAG GAT CTG 3293

17		18
Lys Arg Pro Arg Thr Ala 1035	Arg Arg Tyr Leu A 1040	rg His Leu Glu Asp Leu 1045
GGC TAT GTC CGG CTT AAG	AAG ATC GGC TAC	
Gly Tyr Val Arg Leu Lys		
1050	1055	1060
TCA CTT AAG AAC TAC AGA		
Ser Leu Lys Asn Tyr Arg	Arg Leu Tyr Glu A	la Leu Yal Glu Asn Yal
1065	1070	1075
AGA TAC AAC GGC AAC AAG		
Arg Tyr Asn Gly Asn Lys 1080	1085	al Glu Phe Ash Ser 11e 1090
CGG GAT GCA GTT GGC ATA		
Arg Asp Ala Val Gly Ile		
1095 110	_	105 1110
ATC GGC ACG CTG AAC GGC	TTC AGA ATG AGA A	AG CTC ATT GAA GTG GAC 3533
Ile Gly Thr Leu Asn Gly		
1115	1120	1125
GAG TCG TTA GCA AAG CTC		
Glu Ser Leu Ala Lys Leu		al Ser Glu Gly Tyr Ala
1130	1135	1140
AGA AAG CAG AGG AAT CCC		
Arg Lys Gln Arg Asn Pro 1145	1150	
TAC AAC GAA GAC CCT GAA		1155 TG GAG AGA CTC GCC AGC 3677
Tyr Asn Glu Asp Pro Glu		
1160	1165	1170
AGG TTT TTC GGG AAG GTG	AGG CGG GGC AGG A	AC TAC GTT GAG ATA CCG 3725
Arg Phe Phe Gly Lys Val	Arg Arg Gly Arg A	sn Tyr Val Glu Ile Pro
1175 118	-	185 1190
AAG AAG ATC GGC TAC CTG		
Lys Lys Ile Gly Tyr Leu		
1195 GAG AAC AAG AGG ATT CCC	1200 CAC TTC CTC TTC A	1205
Glu Asn Lys Arg Ile Pro		
1210	1215	1220
CGG CTG GCC TTC CTT GAG	GGG TAC TCA TCG	
Arg Leu Ala Phe Leu Glu		
1225	1230	1235
GAA CAA GAG ACT CAG GCT		
Glu Glo Glo Thr Gln Ala		rg Ala Leu Ala Asn Gln
1240	1245	1250
CTC GTC CTC CTC TTG AAC		
Leu Val Leu Leu Leu Asu 1255 126		
CAC GAC AGC GGC GTT TAC		265 1270 AC CAC CAC CTC CCC TTC 4012
His Asp Ser Gly Val Tyr		
1275	1280	1285
GTA AAG CTG GAC AAG AAA		
Val Lys Leu Asp Lys Lys		

1.	0						``								00	101
AAG GAA		CTG	AGC	GAG	CT(°	777	ccc	AAC	CTT	TTC	CAG	444	440	CTC	20	4109
Lys Gle																4109
2,0 0.0	130	_				1310		D)3	,,,	ı nc	1315	_	11311	741		
AGT CC1			TTC	AGG	AAG			GAG	GAC	GGA			GAT	ccc		4157
Ser Pro																
132	20				1325	5			_	1336)		_			
GAA AAG	GCC	CAG	AGG	CTC	TCC	TGG	CTC	ATT	GAG	GGG	GAC	GTA	GTG	CTC		4205
Glu Lys	Ala	Gln	Arg	Leu	Ser	Trp	Leu	He	Glu	Gly	Asp	Val	Val	Leu		
1335				1340)				134	5				1350)	
GAC CGC																4253
Asp Arg	y Val	Glu			Asp	Val	Glu			Asp	Gly	Tyr				
			135					1360					136			
GAC CTC																4301
Asp Let	ı ser			ASP	ASO	618			rea	184	ыу			Leu		
GTC TAT	CCT	137		ACC	TAC	TAC	137!		TAC	ccr	TAT	1380		CCC		4349
Val Ty																4049
,.	138		,,,,,,,	501	.,.	1390		.,.	131	0.,	139		*111.5	,,,,		
CGC TGC			AAG	GAG	TGT			AGC	GTA	ACG			GGA	AGG		4397
Arg Tr																
140	00				140	5				141	0					
GAG TAG	ATA C	ACG	ATG	ACC	ATC	AAG	GAG	ATA	GAG	GAA	AAG	TAC	GGC	TTT		4445
Glu Ty	I le	Thr	Me t	Thr	Ile	Lys	Glu	He	Glu	Glu	Lys	Tyr	Gly	Phe		
1415				1420	0				142	5				1430)	
AAG GTA			_													4493
Lys Val	lle	Туг			Thr	Asp	Gly			Ala	Thr	He				
CCC CAT		C44	143			140		144			ምም/ [‡]	OTO	144			45.43
GCC GAT																4541
nia noj	AIG	145		741	Lys	LJS	145		MC I	010	1 110	146		1 y 1		
ATC AAG	GCC			CCG	GGC	GCG			CTC	GAG	TAC			TTC		4589
Ile Ası		_	_													
	146					1470					147		·			
TAC AA	A CGC	GGC	TTC	TTC	GTC	ACG	AAG	AAG	AAG	TAT	GCG	GTG	ATA	GAC		4637
Tyr Ly:	Arg	Gly	Phe	Phe	Val	Thr	Lys	Lys	Lys	Tyr	Ala	Val	He	Asp		
148	30				148	5				149	0					
GAG GA																4685
Glu Gli	ı Gly	Lys				Arg	Gly				Val	Arg	Arg			
1495				150					150					1510)	
TGG AGG																4733
Trp Sea	r Glu	11e			GIu	Thr	GIn			Val	Leu	Glu				
CTA AA		የርተ	151		CAC	110	ccc	152		ATA	OTO		152			4701
CTA AAG																4781
ren ra	y vsh	153		741	GIU	L 32	153		VIR	116	Val	154		Yaı		
ACC GAZ	A AAG			AAG	TAC	GAG			CCC	GAG	AAG			ATC		4829
Thr Gl																1023
	154			_,,	- • -	155					155					
CAC GAG			ACG	AGG	GAT			GAC	TAC	AAG			GGT	ccc		1877
His Glo																

```
1560
                                        1565
                                                           1570
                  CAC GTT GCC GTT GCC AAG AGG TTG GCC GCG AGA GGA GTC AAA ATA CGC
                                                                                 4925
                  His Val Ala Val Ala Lys Arg Leu Ala Ala Arg Gly Val Lys Ile Arg
                                    1580
                                                       1585
                  CCT GGA ACG GTG ATA AGC TAC ATC GTG CTC AAG GGC TCT GGG AGG ATA
                                                                                 4973
                  Pro Gly Thr Val Ile Ser Tyr Ile Val Leu Lys Gly Ser Gly Arg Ile
                                 1595
                                                   1600
                  GGC GAC AGG GCG ATA CCG TTC GAC GAG TTC GAC CCG ACG AAG CAC AAG
                                                                                 5023
                  Gly Asp Arg Ala Ile Pro Phe Asp Glu Phe Asp Pro Thr Lys His Lys
                                               1615
                  TAC GAC GCC GAG TAC TAC ATT GAG AAC CAG GTT CTC CCA GCC GTT GAG
                                                                                 5069
                  Tyr Asp Ala Glu Tyr Tyr Ile Glu Asn Glo Val Leu Pro Ala Val Glu
                                            1630
                                                              1635
                  AGA ATT CTG AGA GCC TTC GGT TAC CGC AAG GAA GAC CTG CGC TAC CAG
                                                                                 5117
                  Arg lie Leu Arg Ala Phe Gly Tyr Arg Lys Glu Asp Leu Arg Tyr Gln
                                        1645
                                                           1650
                  AAG ACG AGA CAG GTT GGT TTG AGT GCT TGG CTG AAG CCG AAG GGA ACT
                                                                                 5165
                  Lys Thr Arg Gln Val Gly Leu Ser Ala Trp Leu Lys Pro Lys Gly Thr
                  1655
                                    1660
                                                       1665
                  TGACCTTTCC ATTTGTTTC CAGCGGATAA CCCTTTAACT TCCCTTTCAA AAACTCCCT
                                                                                 5225
                  TAGGGAAAGA CCATGAAGAT AGAAATCCGG CGGCGCCCGG TTAAATACGC TAGGATAGA
                  GTGAAGCCAG ACGGCAGGGT AGTCGTCACT GCCCCGAGGG TTCAACGTTG AGAAGTT
                                                                                 5342
【0032】配列番号2
                                                       トポロジー:直鎖状
配列の長さ:774
                                                      配列の種類:タンパク質
配列の型:アミノ酸
                  Met Ile Leu Asp Thr Asp Tyr Ile Thr Glu Asp Gly Lys Pro Val Ile
                                 5
                                                    10
                  Arg Ile Phe Lys Lys Glu Asn Gly Glu Phe Lys Ile Glu Tyr Asp Arg
                             20
                                             25
                  Thr Phe Glu Pro Tyr Phe Tyr Ala Leu Leu Lys Asp Asp Ser Ala Ile
                          35 40
                                                              45
                  Glu Glu Val Lys Lys lie Thr Ala Glu Arg His Gly Thr Val Val Thr
                       50 55
                                                60
                  Val Lys Arg Val Glu Lys Val Gln Lys Lys Phe Leu Gly Arg Pro Val
                   65 70
                                                      75
                  Glu Val Trp Lys Leu Tyr Phe Thr His Pro Gln Asp Val Pro Ala Ile
                                 85
                                                   90
                  Arg Asp Lys Ile Arg Glu His Gly Ala Val Ile Asp Ile Tyr Glu Tyr
                                    i,
                                             105
                  Asp lie Pro Phe Ala Lys Arg Tyr Leu lie Asp Lys Gly Leu Val Pro
                                         120
                  Met Glu Gly Asp Glu Glu Leu Lys Met Leu Ala Phe Asp Ile Gln Thr
                                       135
                                                          140
                  Leu Tyr His Glu Gly Glu Glu Phe Ala Glu Gly Pro Ile Leu Met Ile
                                   150
                                                     155
                  Ser Tyr Ala Asp Glu Glu Gly Ala Arg Val Ile Thr Trp Lys Asn Val
                                165
                                                  170
                  Asp Leu Pro Tyr Val Asp Val Val Ser Thr Glu Arg Glu Met Ile Lys
```

180

Arg Phe Leu Arg Val Val Lys Glu Lys Asp Pro Asp Val Leu Ile Thr

200 Tyr Asn Gly Asp Asn Phe Asp Phe Ala Tyr Leu Lys Lys Arg Cys Glu 215

220 Lys Leu Gly Ile Asn Phe Ala Leu Gly Arg Asp Gly Ser Glu Pro Lys

230 235 Ile Gin Arg Met Giy Asp Arg Phe Ala Val Giu Val Lys Giy Arg Ile

245 250 His Phe Asp Leu Tyr Pro Val Ile Arg Arg Thr Ile Asn Leu Pro Thr

265 Tyr Thr Leu Glu Ala Val Tyr Glu Ala Val Phe Gly Glu Pro Lys Glu

280 Lys Val Tyr Ala Glu Glu Ile Thr Pro Ala Trp Glu Thr Gly Glu Asn

295 300 Leu Glu Arg Val Ala Arg Tyr Ser Met Glu Asp Ala Lys Val Thr Tyr

310 315 Glu Leu Gly Lys Glu Phe Leu Pro Met Glu Ala Gln Leu Ser Arg Leu

330

325

lle Gly Gin Ser Leu Trp Asp Val Ser Arg Ser Ser Thr Gly Asn Leu 345

Val Glu Trp Phe Leu Leu Arg Lys Ala Tyr Glu Arg Asn Glu Leu Ala 360 365

Pro Asn Lys Pro Asp Glu Lys Glu Leu Ala Arg Arg Arg Gin Ser Tyr 375

Glu Gly Gly Tyr Vai Lys Glu Pro Glu Arg Gly Leu Trp Glu Asn Ile 390 395

Val Tyr Leu Asp Phe Arg Ser Leu Tyr Pro Ser Ile Ile Ile Thr His 410

Asn Val Ser Pro Asp Thr Leu Asn Arg Glu Gly Cys Lys Glu Tyr Asp 420 425

Val Ala Pro Gln Val Gly His Arg Phe Cys Lys Asp Phe Pro Gly Phe 440

lle Pro Ser Leu Leu Gly Asp Leu Leu Glu Glu Arg Gln Lys 11e Lys 455 460

Lys Lys Met Lys Ala Thr Ile Asp Pro Ile Glu Arg Lys Leu Leu Asp 470 475

Tyr Arg Glm Arg Ala Ile Lys Ile Leu Ala Asm Ser Tyr Tyr Gly Tyr 485 490

Tyr Gly Tyr Ala Arg Ala Arg Trp Tyr Cys Lys Glu Cys Ala Glu Ser 505

Val Thr Ala Trp Gly Arg Glu Tyr Ile Thr Met Thr Ile Lys Glu Ile 520

Glu Glu Lys Tyr Gly Phe Lys Val Ile Tyr Ser Asp Thr Asp Gly Phe 535 540

Phe Ala Thr lle Pro Gly Ala Asp Ala Glu Thr Val Lys Lys Ala 550 555

Met Glu Phe Leu Asn Tyr Ile Asn Ala Lys Leu Pro Gly Ala Leu Glu

565 570

Leu Glu Tyr Glu Gly Phe Tyr Lys Arg Gly Phe Phe Val Thr Lys Lys 585

```
Lys Tyr Ala Val Ile Asp Glu Glu Gly Lys Ile Thr Thr Arg Gly Leu
        595
                           600
Glu Ile Val Arg Arg Asp Trp Ser Glu Ile Ala Lys Glu Thr Gln Ala
                       615
                                           620
Arg Val Leu Glu Ala Leu Leu Lys Asp Gly Asp Val Glu Lys Ala Val
625
                    630
                                       635
Arg Ile Val Lys Glu Val Thr Glu Lys Leu Ser Lys Tyr Glu Val Pro
               645
                                   650
Pro Glu Lys Leu Val Ile His Glu Gln Ile Thr Arg Asp Leu Lys Asp
                               665
                                                   670
Tyr Lys Ala Thr Gly Pro His Val Ala Val Ala Lys Arg Leu Ala Ala
        675
                           680
Arg Gly Val Lys Ile Arg Pro Gly Thr Val Ile Ser Tyr Ile Val Leu
                       695
                                           700
Lys Gly Ser Gly Arg Ile Gly Asp Arg Ala Ile Pro Phe Asp Glu Phe
                   710
                                      715
Asp Pro Thr Lys His Lys Tyr Asp Ala Glu Tyr Tyr Ile Glu Asn Gln
                                   730
Val Leu Pro Ala Val Glu Arg Ile Leu Arg Ala Phe Gly Tyr Arg Lys
           740
                               745
Glu Asp Leu Arg Tyr Gln Lys Thr Arg Gln Val Gly Leu Ser Ala Trp
       755
                           760
                                               765
Leu Lys Pro Lys Gly Thr
```

[0033]配列番号3

配列の長さ:5342 配列の型:核酸(DNA)

鎖の数:2本鎖

トロポジー:直鎖状

配列の種類: c DNA 起源:超好熱始原菌

株名:KOD1

配列

770

GCTTGAGGGC CTGCGGTTAT GGGACGTTGC AGTTTGCGCC TACTCAAAGA TGCCGGTTTT ATAACGGAGA AAAATGGGGA GCTATTACGA TCTCTCCTTG ATGTGGGGTT TACAATAAAG 120 CCTGGATTGT TCTACAAGAT TATGGGGGAT GAAAGATGAT CCTCGACACT GACTACATAA 180 CCGAGGATGG AAAGCCTGTC ATAAGAATTT TCAAGAAGGA AAACGGCGAG TTTAAGATTG AGTACGACCG GACTTTTGAA CCCTACTTCT ACGCCCTCCT GAAGGACGAT TCTGCCATTG AGGAAGTCAA GAAGATAACC GCCGAGAGGC ACGGGACGGT TGTAACGGTT AAGCGGGTTG 360 AAAAGGTTCA GAAGAAGTTC CTCGGGAGAC CAGTTGAGGT CTGGAAACTC TACTTTACTC 420 ATCCGCAGGA CGTCCCAGCG ATAAGGGACA AGATACGAGA GCATGGAGCA GTTATTGACA 480 TCTACGAGTA CGACATACCC TTCGCCAAGC GCTACCTCAT AGACAAGGGA TTAGTGCCAA 540 TGGAAGGCGA CGAGGAGCTG AAAATGCTCG CCTTCGACAT TCAAACTCTC TACCATGAGG 600 GCGAGGAGTT CGCCGAGGGG CCAATCCTTA TGATAAGCTA CGCCGACGAG GAAGGGGCCA 660 GGGTGATAAC TTGGAAGAAC GTGGATCTCC CCTACGTTGA CGTCGTCTCG ACGGAGAGGG 720 AGATGATAAA GCGCTTCCTC CGTGTTGTGA AGGAGAAAGA CCCGGACGTT CTCATAACCT ACAACGGCGA CAACTTCGAC TTCGCCTATC TGAAAAAGCG CTGTGAAAAG CTCGGAATAA 840 ACTTCGCCCT CGGAAGGGAT GGAAGCGAGC CGAAGATTCA GAGGATGGGC GACAGGTTTG 900 CCGTCGAAGT GAAGGGACGG ATACACTTCG ATCTCTATCC TGTGATAAGA CGGACGATAA 960 ACCTGCCCAC ATACACGCTT GAGGCCGTTT ATGAAGCCGT CTTCGGTCAG CCGAAGGAGA 1020 AGGTTTACGC TGAGGAAATA ACACCAGCCT GGGAAACCGG CGAGAACCTT GAGAGAGTCG 1080 CCCGCTACTC GATGGAAGAT GCGAAGGTCA CATACGAGCT TGGGAAGGAG TTCCTTCCGA 1140 TGGAGGCCCA GCTTTCTCGC TTAATCGGCC AGTCCCTCTG GGACGTCTCC CGCTCCAGCA 1200 CTGGCAACCT CGTTGAGTGG TTCCTCCTCA GGAAGGCCCT ATGAGAGGAA TGAGCTGGCC 1260

28

CCGAACAAGC CCGATGAAAA GGAGCTGGCC AGAAGACGGC AGAGCTATGA AGGAGGCTAT 1320

GTAAAAGAGC CCGAGAGAGG GTTGTGGGAG AACATAGTGT ACCTAGATTT TAGATGCCAT 1380

CCAGCCGATA CGAAGGTTGT CGTCAAGGGG AAGGGGATTA TAAACATCAG CGAGGTTCAG 1440

GAAGGTGACT ATGTCCTTGG GATTGACGGC TGGCAGAGAG TTAGAAAAGT ATGGGAATAC 1500

GACTACAAAG GGGAGCTTGT AAACATAAAC GGGTTAAAGT GTACGCCCAA TCATAAGCTT 1560

CCCGTTGTTA CAAAGAACGA ACGACAAACG AGAATAAGAG ACAGTCTTGC TAAGTCTTTC 1620

CTTACTAAAA AAGTTAAGGG CAAGATAATA ACCACTCCCC TTTTCTATGA AATAGGCAGA, 1680

GCGACAAGTG AGAATATTCC AGAAGAAGAG GTTCTCAAGG GAGAGCTCGC TGGCATAGTA, 1740 TTGGCTGAAG GAACGCTCTT GAGGAAAGAC GTTGAATACT TTGATTCATC CCGCAAAAAA 1800

CGGAGGATTT CACACCAGTA TCGTGTTGAG ATAACCATTG GGAAAGACGA GGAGGAGTTT 1860

AGGGATCGTA TCACATACAT TTTTGAGCGT TTGTTTGGGA TTACTCCAAG CATCTCGGAG 1920

AAGAAAGGAA CTAACGCAGT AACACTCAAA GTTGCGAAGA AGAATGTTTA TCTTAAAGTC 1980

AAGGAAATTA TGGACAACAT AGAGTCCCTA CATGCCCCCT CGGTTCTCAG GGGATTCTTC 2040

GAAGGCGACG GTTCAGTAAA CAGGTTAGGA GGAGTATTGT TGCAACCCAG GGTACAAAGA 2100

ACGACTGGAA GATTAAACTG GTGTCAAAAC TGCTCTCCCA GCTTGGTATC CCTCATCAAA 2160

CCTACACGTA TCAGTATCAG GAAAATGGGA AAGATCGGAG CAGGTATATA CTGGAGATAA 2220

CTGGAAAGGA CGGATTGATA CTGTTCCAAA CACTCATTGG ATTCATCAGT GAAAGAAAGA 2280 ACGCTCTGCT TAATAAGGCA ATATCTCAGA GGGAAATGAA CAACTTGGAA AACAATGGAT 2340

TTTACAGGCT CAGTGAATTC AATGTCAGCA CGGAATACTA TGAGGGCAAG GTCTATGACT 2400

TAACTCTTGA AGGAACTCCC TACTTTGCCA ATGGCATATT GACCCATAAC TCCCTGTACC 2460

CCTCAATCAT CATCACCCAC AACGTCTCGC CGGATACGCT CAACAGAGAA GGATGCAAGG 2520

AATATGACGT TGCCCCACAG GTCGGCCACC GCTTCTGCAA GGACTTCCCA GGATTTATCC 2580

CGAGCCTGCT TGGAGACCTC CTAGAGGAGA GGCAGAAGAT AAAGAAGAAG ATGAAGGCCA 2640

CGATTGACCC GATCGAGAGG AAGCTCCTCG ATTACAGGCA GAGGGCCATC AAGATCCTGG 2700 CAAACAGCAT CCTACCGAG GAATGGCTTC CAGTCCTCGA GGAAGGGGAG GTTCACTTCG 2760

TCAGGATTGG AGAGCTCATA GACCGGATGA TGGAGGAAAA TGCTGGGAAA GTAAAGAGAG 2820

AGGGCGAGAC GGAAGTGCTT GAGGTCAGTG GGCTTGAAGT CCCGTCCTTT AACAGGAGAA 2880

CTAACAAGGC CGAGCTCAAG AGAGTAAAGG CCCTGATTAG GCACGATTAT TCTGGCAAGG 2940

TCTACACCAT CAGACTGAAG TCGGGGAGGA GAATAAAGAT AACCTCTGGC CACAGCCTCT 3000

TCTCTGTGAG AAACGGGGAG CTCGTTGAAG TTACGGGCGA TGAACTAAAG CCAGGTGACC 3060

TCGTTGCAGT CCCGCGGAGA TTGGAGCTTC CTGAGAGAAA CCACGTGCTG AACCTCGTTG 3120

AACTGCTCCT TGGAACGCCA GAAGAAGAAA CTTTGGACAT CGTCATGACG ATCCCAGTCA 3180

AGGGTAAGAA GAACTTCTTT AAAGGGATGC TCAGGACTTT GCGCTGGATT TTCGGAGAGG 3240

AAAAGAGGCC CAGAACCGCG AGACGCTATC TCAGGCACCT TGAGGATCTG GGCTATGTCC 3300 GGCTTAAGAA GATCGGCTAC GAAGTCCTCG ACTGGGACTC ACTTAAGAAC TACAGAAGGC 3360

TCTACGAGGC GCTTGTCGAG AACGTCAGAT ACAACGGCAA CAAGAGGGAG TACCTCGTTG 3420

AATTCAATTC CATCCGGGAT GCAGTTGGCA TAATGCCCCT AAAAGAGCTG AAGGAGTGGA 3480

AGATCGGCAC GCTGAACGGC TTCAGAATGA GAAAGCTCAT TGAAGTGGAC GAGTCGTTAG 3540

CAAAGCTCCT CGGCTACTAC GTGAGCGAGG GCTATGCAAG AAAGCAGAGG AATCCCAAAA 3600

ACGGCTGGAG CTACAGCGTG AAGCTCTACA ACGAAGACCC TGAAGTGCTG GACGATATGG 3660

AGAGACTCGC CAGCAGGTTT TTCGGGAAGG TGAGGCGGGG CAGGAACTAC GTTGAGATAC 3720

CGAAGAAGAT CGGCTACCTG CTCTTTGAGA ACATGTGCGG TGTCCTAGCG GAGAACAAGA 3780

GGATTCCCGA GTTCGTCTTC ACGTCCCCGA AAGGGGTTCG, GCTGGCCTTC CTTGAGGGGT 3840

ACTCATCGGC GATGGCGACG TCCACCGAAC AAGAGACTCA GGCTCTCAAC GAAAAGCGAG 3900

CTTTAGCGAA CCAGCTCGTC CTCCTCTTGA ACTCGGTGGG GGTCTCTGCT GTAAAACTTG 3960 GGCACGACAG CGGCGTTTAC AGGGTCTATA TAAACGAGGA GCTCCCGTTC GTAAAGCTGG 4020

ACAAGAAAAA GAACGCCTAC TACTCACACG TGATCCCCAA GGAAGTCCTG AGCGAGGTCT 4080

TTGGGAAGGT TTTCCAGAAA AACGTCAGTC CTCAGACCTT CAGGAAGATG GTCGAGGACG 4140

GAAGACTCGA TCCCGAAAAG GCCCAGAGGC TCTCCTGGCT CATTGAGGGG GACGTAGTGC 4200

TCGACCGCGT TGAGTCCGTT GATGTGGAAG ACTACGATGG TTATGTCTAT GACCTGAGCG 4260

```
TCGAGGACAA CGAGAACTTC CTCGTTGGCT TTGGGTTGGT CTATGCTCAC AACAGCTACT 4320
                  ACGGTTACTA CGGCTATGCA AGGGCGCGCT GGTACTGCAA GGAGTGTGCA GAGAGCGTAA 4380
                  CGGCCTGGGG AAGGGAGTAC ATAACGATGA CCATCAAGGA GATAGAGGAA AAGTACGGCT 4440
                  TTAAGGTAAT CTACAGCGAC ACCGACGGAT TTTTTGCCAC AATACCTGGA GCCGATGCTG 4500
                  AAACCGTCAA AAAGAAGGCT ATGGAGTTCC TCAACTATAT CAACGCCAAA CTTCCGGGCG 4560
                  CGCTTGAGCT CGAGTACGAG GGCTTCTACA AACGCGGCTT CTTCGTCACG AAGAAGAAGT 4620
                  ATGCGGTGAT AGACGAGGAA GGCAAGATAA CAACGCGCGG ACTTGAGATT GTGAGGCGTG 4680
                  ACTGGAGCGA GATAGCGAAA GAGACGCAGG CGAGGGTTCT TGAAGCTTTG CTAAAGGACG 1710
                  GTGACGTCGA GAAGGCCGTG AGGATAGTCA AAGAAGTTAC CGAAAAGCTG AGCAAGTACG 4800
                  AGGTTCCGCC GGAGAAGCTG GTGATCCACG AGCAGATAAC GAGGGATTTA AAGGACTACA 4860
                  AGGCAACCGG TCCCCACGTT GCCGTTGCCA AGAGGTTGGC CGCGAGAGGA GTCAAAATAC 4920
                  GCCCTGGAAC GGTGATAAGC TACATCGTGC TCAAGGGCTC TGGGAGGATA GGCGACAGGG 4980
                  CGATACCGTT CGACGAGTTC GACCCGACGA AGCACAAGTA CGATGCCGAG TACTACATTG 5040
                  AGAACCAGGT TCTCCCAGCC GTTGAGAGAA TTCTGAGAGC CTTCGGTTAC CGCAAGGAAG 5100
                 ACCTGCGCTA CCAGAAGACG AGACAGGTTG GTTTGAGTGC TTGGCTGAAG CCGAAGGGAA 5160
                 CTTGACCTTT CCATTTGTTT TCCAGCGGAT AACCCTTTAA CTTCCCTTTC AAAAACTCCC 5220
                 TTTAGGGAAA GACCATGAAG ATAGAAATCC GGCGGCGCCC GGTTAAATAC GCTAGGATAG 5280
                 AAGTGAAGCC AGACGGCAGG GTAGTCGTCA CTGCCCCGAG GGTTCAACGT TGAGAAGTT 5339
 【0034】配列番号4
                                                    *トポロジー:直鎖状
配列の長さ:24
                                                  20 配列の種類: 合成DNA
配列の型:核酸
                  配列
                 GGATTAGTGC CAATGGAAGG CGAC
 【0035】配列番号5
                                                    ※トポロジー:直鎖状
配列の長さ:24
                                                      配列の種類:合成DNA
配列の型:核酸
                                                ×
                  配列
                 GAGGGCGAAG TITATTCCGA GCTT
                                                                   24
【0036】配列番号6
                                                    ★鎖の数:2本鎖
配列の長さ:324
                                                    トロポジー:直鎖状
配列の型:核酸(DNA)
                                                      配列の種類:cDNA
                  配列
                  GGATTAGTGC CAATGGAAGG CGACGAGGAG CTGAAAATGC TCGCCTTCGA CATTCAAACT
                 CTCTACCATG AGGGCGAGGA GTTCGCCGAG GGGCCAATCC TTATGATAAG CTACGCCGAC 120
                  GAGGAAGGGG CCAGGGTGAT AACTTGGAAG AACGTGGATC TCCCCTACGT TGACGTCGTC 180
                  TCGACGGAGA GGGAGATGAT AAAGCGCTTC CTCCGTGTTG TGAAGGAGAA AGACCCGGAC 240
                  GTTCTCATAA CCTACAACGG CGACAACTTC GACTTCGCCT ATCTGAAAAA GCGCTGTGAA 300
                  AAGCTCGGAA TAAACTTCGC CCTC
                                                                               324
 【0037】配列番号7
                                                      トポロジー:直鎖状
配列の長さ:108
                                                  40 配列の種類:タンパク質
配列の型:アミノ酸
                  配列
                  Gly Leu Val Pro Met Glu Gly Asp Glu Glu Leu Lys Met Leu Ala Phe
                                  5
                                                   10
                  Asp Ile Gln Thr Leu Tyr His Glu Gly Glu Glu Phe Ala Glu Gly Pro
                                                25
                  Ile Leu Met Ile Ser Tyr Ala Asp Glu Glu Gly Ala Arg Val Ile Thr
                                            40
                  Trp Lys Asn Val Asp Leu Pro Tyr Val Asp Val Val Ser Thr Glu Arg
```

55

Glu Met Ile Lys Arg Phe Leu Arg Vai Val Lys Glu Lys Asp Pro Asp

70 80

105

Val Leu Ile Thr Tyr Asn Gly Asp Asn Phe Asp Phe Ala Tyr Leu Lys 85 90

Lys Arg Cys Glu Lys Leu Gly Ile Asn Phe Ala Leu

100 -

【0038】配列番号8

*鎖の数: 1本鎖

配列の長さ:42

配列の種類:合成DNA

配列の型:核酸(DNA)

GCCATCAAGA TCCTGGCAAA CAGCTACTAC GGTTACTACG GC

32

【0039】配列番号9

※鎖の数: 1本鎖

配列の種類:合成DNA

配列の長さ:32

配列の型:核酸(DNA)

GATGGATCCA ACTTCTCAAC GTTGAACCCT CG

32

【0040】配列番号10

★鎖の数: 1本鎖

配列の種類:合成DNA

配列の長さ:46 配列の型:核酸(DNA)

配列

GAACATAGTG TACCTAGATT TTAGATCCCT GTACCCCTCA ATCATC

【0041】配列番号11

☆鎖の数: 1本鎖

配列の長さ:42

配列の種類:合成DNA

配列の型:核酸 (DNA)

GCCGTAGTAA CCGTAGTAGC TGTTTGCCAG GATCTTGATG GC

【0042】配列番号12

◆鎖の数: 1本鎖 配列の種類:合成DNA

配列の長さ:33 配列の型:核酸(DNA)

来の耐熱性DNAポリメラーゼ遺伝子およびThermococc

40 us litoralis由来の耐熱性DNAポリメラーゼ遺伝子と

ATCGATATCC TCGACACTGA CTACATAACC GAG

【0043】配列番号13

*鎖の数: 1本鎖

配列の長さ:46

配列の種類:合成DNA

配列の型:核酸(DNA)

配列

GATGATTGAG GGGTACAGGG ATCTAAAATC TAGGTACACT ATGTTC

【図面の簡単な説明】

【図4】 超好熱始原菌KOD1由来のDNAポリメラ - ゼ遺伝子と類縁菌と思われる Pyrococcus furiosus由

【図1】 組換え発現ペクターの構築図を示す。 【図2】 KOD1由来耐熱性DNAポリメラーゼ分子

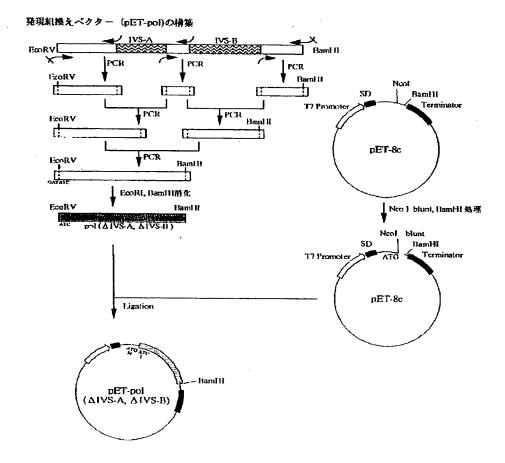
量測定結果を示す電気泳動の写真である。

【図3】 KOD1由来耐熱性DNAポリメラーゼによ

るPCRの結果を示す電気泳動の写真である。

の比較を示す。

[図1]

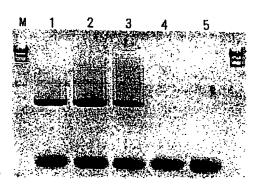


【図2】

1 2 3 4 5 6 M 27 1 V 2 V 4 X Y 43 k

- 1: pET-8c 沈殿
- 2: pET-pol(AIVS-A, AIVS-B) 沈殿
- 3: pET-8c 上澄み
- 4: pET-8c 上澄み x5
- 5: pET-po!(&IVS-A, &IVS-B) 上澄み
- 6: pET-pol(AIVS-AAIVS-B) 上澄み x5

組換え歯が生産する超好熱始原菌KOD1株由来 DNAポリメラーゼの分子量測定 (SDS-PAGE法) [図3]

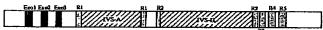


- 1:Vent ポリメラーゼ (Thermococcus litoralis由来)
- 2:pET-pol(△IVS-A, △IVS-B)上澄み
- 3: pET-pol(△IVS-A, △IVS-B)上澄みx5
- 4:pET-8c 上澄み
- 5:pET-Bc 上澄み x 5

組換え猫が生産する超好熱蛤原菌KOD1株由来 DNAポリメラーゼを用いたPCR(Polymerase Chain Reaction)により増幅されたDNA断片

【図4】

超行熱始原薬KOD1株のDNAポリメラーゼ遺伝子



Pyrococcus furiosusのDNAポリメラーゼ遺伝子 (Pfu DNA polymerase)



Thermococcus literalisのDNAポリメラーゼ遺伝子 (Vent DNA polymerase)



超好熱站原稿KODI株のDNAポリメラーゼ遺伝子と他の好熱性菌の DNAポリメラーゼ遺伝子の比較

フロントページの続き

(51) Int. Cl. 5

識別記号 庁内整理番号

技術表示箇所

C 1 2 R 1:19) (C 1 2 N 1/21 FΙ

C 1 2 R 1:19)